Molecular Ecology Notes (2005) **5**, 934–937 doi: 10.1111/j.1471-8286.2005.01121.x

PRIMER NOTE

Characterization and evaluation of microsatellite loci in European hazelnut (*Corylus avellana* L.) and their transferability to other *Corylus* species

P. BOCCACCI,* A. AKKAK,* N. V. BASSIL,† S. A. MEHLENBACHER‡ and R. BOTTA*
**Dipartimento di Colture Arboree, Università di Torino, Via Leonardo da Vinci 44, 10095 Grugliasco (TO), Italy, †USDA-ARS National Clonal Germplasm Repository, 33447 Peoria Rd., Corvallis, OR 97331 USA, ‡Department of Horticulture, Oregon State University, 4017 Agricultural and Life Sciences Bldg., Corvallis, OR 97331 USA*

Abstract

In this work, 18 microsatellite loci were developed in the European hazelnut (*Corylus avellana* L.) using three enriched genomic libraries. They were evaluated on a set of 20 accessions of this species on the basis of number of alleles (mean: 7.1), expected heterozygosity (mean: 0.67), power of discrimination (mean: 0.77) and polymorphism information content (mean: 0.64). Cross-species transferability was evaluated using seven other *Corylus* species. All primer pairs amplified in all species, except for CaT-C505 in *Corylus ferox* and CaT-A114 in *Corylus californica*.

Keywords: cross-species amplification, filbert, polymorphism, simple sequence repeats

Received 05 May 2005; revision accepted 22 June 2005

Corylus species are widely distributed throughout temperate regions of the Northern Hemisphere from Japan, Korea, China and the Russian Far East to the Caucasus, Turkey, Europe and North America (Kasapligil 1972).

The European hazelnut (*Corylus avellana* L.) is a commercially important species. Cultivars in Europe and Turkey were selected, over many centuries, from local wild populations. In addition, several wild species have been crossed with the economically important *C. avellana* in an effort to obtain specific traits (Mehlenbacher 1991). Identification of hazelnut cultivars is primarily based on analysis of nuts, husks and other morphological traits. These, however, are often unreliable or imprecise indicators of plant genotype, being influenced by environmental factors. Thus, discrimination among closely related cultivars and clones is often extremely difficult. Cultivar identification is especially difficult in a nursery setting where plants are young and not yet bearing.

DNA typing can be a convenient method for accurately identifying hazelnut cultivars. Microsatellite or simple sequence repeat (SSR) molecular markers have been routinely isolated from plants and have been very useful because they are locus-specific, codominant, highly polymorphic and highly reproducible.

Correspondence: R. Botta, Fax: +39-011-6708658; E-mail: roberto.botta@unito.it

In this work, 18 microsatellite loci were developed in *C. avellana* and their polymorphism was studied in a set of 20 accessions of this species, coming from the various countries that grow hazelnut (Table 1). Their cross-species transferability was evaluated using representatives of seven other *Corylus* species (Table 1): *C. colurna* L., *C. californica* Marshall, *C. ferox* Wallich, *C. heterophylla* Fischer, *C. papyracea* Hickel, *C. chinensis* Franchet and *C. americana* Marshall. Furthermore, five individuals from the cross 'Tonda Gentile delle Langhe' × 'Cosford' (Romisondo *et al*. 1983) and the parent cultivars were analysed to check the segregation of alleles at each locus.

Genetic Identification Services in Chatsworth, California, USA constructed three genomic libraries enriched for CA-, GA- and GAA-repeats, respectively, for Oregon State University (OSU), Corvallis, Oregon, USA, using DNA from dark-germinated seeds of a mixture of *C. avellana* cultivars and selections. Later, DNA was cloned into pUC19 vector and used to transform *Escherichia coli* DH5α following the method of Edwards *et al*. (1996). DNA sequencing was performed with a BigDye Terminator Cycle Sequencing Ready Reaction kit version 2.0 using an ABI PRISM 377 sequencer (Applied Biosystems). All sequences obtained were blasted against each other using the BIOEDIT version 5.0.9 package (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) and duplicate sequences were removed. Primers were

Table 1 Genotypes used in this study. Twenty *Corylus avellana* genotypes were used to characterize microsatellite markers, and representatives of seven *Corylus* species were used to investigate cross-species transferability

Name	Accession no.	Species	Origin
252.146	None	*C. avellana* L.	Oregon (USA)
414.062	None	*C. avellana* L.	Oregon (USA)
681.078	CCOR780	*C. avellana* L.	Russia
B-4 Pelargonia	PI 557125	*C. avellana* L.	Macedonia
Bulgaria XI-8	PI 557219	*C. avellana* L.	Bulgaria
Casina	PI 557033	*C. avellana* L.	Spain
Contorta	PI 557049	*C. avellana* L.	UK
Cosford	PI 557039	*C. avellana* L.	UK
Da Viega	PI 557224	*C. avellana* L.	Portugal
Ganja	CCOR771	*C. avellana* L.	Georgia
Gasaway	PI 557042	*C. avellana* L.	USA
Giresun 54.021	PI 557057	*C. avellana* L.	USA
Hall's Giant	PI 557027	*C. avellana* L.	Germany
Imperial de Trebizonde	PI 271105	*C. avellana* L.	Turkey
Negret	PI 270340	*C. avellana* L.	Spain
Pellicule rouge	PI 271110	*C. avellana* L. (*C. maxima* Miller)	France
Romai	PI 557233	*C. avellana* L.	Italy
San Giovanni	PI 557117	*C. avellana* L.	Italy
Tonda Gentile delle Langhe	PI 557035	*C. avellana* L.	Italy
Tonda Gentile Romana	PI 557025	*C. avellana* L.	Italy
640.001	PI 557251	*C. colurna* L.	—
B0509	None	*C. californica*	USA
C. ferox	PI 557302	*C. ferox* Wallich	China
Ogyoo	PI 557323	*C. heterophylla* Fischer	South Korea
Paperbark	None	*C. papyraceae* Hickel	Southern China
W-5	CCOR698	*C. chinensis* Franchet	Southern China
Winkler	PI 557019	*C. americana* Marshall	USA
X-11	PI 557249	*C. colurna* L.	USA

designed using the software PRIMER EXPRESS 1.0 (Applied Biosystems). Polymerase chain reaction (PCR) was performed in a volume of 20 μL containing 50 ng DNA, 0.5 U Ampli*Taq* Gold polymerase (Applied Biosystems), 2 μL 10 × PCR buffer (100 mM Tris-HCl, pH 8.3, 500 mM KCl), 2 mM $MgCl_2$, 200 μM dNTPs and 0.5 μM of each primer. Reactions were carried out under the following conditions: 95 °C for 9 min, 28 cycles of 30 s at 95 °C, 45 s at the annealing temperature for each primer pair (Table 2), and 90 s at 72 °C with a final elongation step of 72 °C for 45 min. Forward primers were labelled with a fluorochrome (6-FAM, HEX, or NED) and PCR products were analysed using an ABI PRISM 377 machine. Allele sizes were determined with GENESCAN software (Applied Biosystems) and data were analysed using the software IDENTITY 1.0 (Wagner & Sefc 1999) and POWER MARKER version 3.02 (http://www.powermarker.net). Power of discrimination was calculated as $PD = 1 - \Sigma g_i^2$, where g_i is the frequency of the ith genotype (Kloosterman *et al*. 1993), and null allele frequency as $r = (H_E - H_O)/(1 + H_E)$, where H_E = expected heterozygosity and H_O = observed heterozygosity (Brookfield 1996). Polymorphism information content (PIC) was determined as $PIC = 1 - \Sigma p_i^2 - \Sigma p_i^2 p_j^2$ (Botstein *et al*. 1980), where p_i and p_j are the frequencies of alleles ith and jth, respectively.

In the small progeny of five seedlings, alleles at all loci appeared to segregate according to Mendelian expectations. Fourteen dinucleotide- and four trinucleotide-containing microsatellite loci were analysed using DNA extracted from the 20 genotypes of *C. avellana* and from the seven *Corylus* species. The polymorphism and discriminant power of each locus (Table 2) were evaluated on the basis of the number of alleles (mean: 7.1), expected heterozygosity (mean: 0.67), power of discrimination (mean: 0.77) and PIC (mean: 0.64).

CaT-B501, CaT-B502, CaT-B503, CaT-B504, CaT-B507, CaT-B508, CaT-B511, CaT-B106 and CaT-B107 are the most interesting loci for fingerprinting on the basis of their high level of polymorphism. Moreover, CaT-A114, CaT-B505, CaT-B509, CaT-C001 and CaT-C504 can be considered promising, in spite of a lower number of alleles; their PD and H_E values are above average and thus a high level of polymorphism is expected once a larger number of genotypes is analysed. Less polymorphic loci (CaT-C502, CaT-C503, CaT-C505 and CaT-B512) might be useful for genetic mapping and other purposes.

Table 2 Repeat motifs, primer sequences, annealing temperature (T_a), allele size ranges (in bp), number of alleles, expected heterozygosity (H_E), observed heterozygosity (H_O), frequency of null alleles (r), polymorphism information content (PIC) and power of discrimination (PD) of 18 SSR loci analysed in 20 *Corylus avellana* cultivars

Locus	Repeat type	GenBank Accession no.	Primer sequences (5′–3′)	T_a (°C)	Allele size range (bp)	No. alleles	H_E	H_O	r	PIC	PD
CaT-A114	Compound $(TA)_2(TG)_{17}(TA)_3$	DQ090024	F: CGCCTTGATAGTATGTTCAAAC R: CGGCAGAATGTAGAAGTCCCC	55	165–181	7	0.80	0.90	−0.058	0.77	0.89
CaT-B106	Imperfect $(AG)_{17}AA(AG)_6$	DQ090007	F: CCAATCGCCAATGAATCATC R: CCCTTTCCAAACTGGGCAT	55	156–181	10	0.85	0.85	−0.001	0.83	0.93
CaT-B107	Perfect $(CT)_{14}$	DQ090008	F: GTAGGTGCACTTGATGTGCTTTAC R: AACACCATATTGAGTCTTTCAAAGC	55	114–153	10	0.82	0.90	−0.046	0.79	0.89
CaT-B501	Perfect $(GA)_{21}$	DQ090009	F: GAAATTCAATCACACCAATAAAGCA R: CCTCCCTTGTCCTCATCACTG	55	117–137	8	0.75	0.70	0.029	0.72	0.89
CaT-B502	Imperfect $(CT)_{16}GCTTTTC(CT)_5$	DQ090010	F: CTCATGACTGCCCATTTCTCG R: AGGCATGCAGGCTTCACAC	50	185–214	9	0.76	0.80	−0.023	0.73	0.91
CaT-B503	Perfect $(GA)_{18}$	DQ090011	F: CTCAATTCACTCGAACGGATAC R: AGCCGATACCAGCCTCTCGC	55	117–133	8	0.81	0.85	−0.020	0.79	0.90
CaT-B504	Perfect $(CT)_{18}$	DQ090012	F: CGCCATCTCCATTTCCCAAC R: CGGAATGGTTTTCTGCTTCAG	55	161–187	9	0.83	0.75	0.045	0.81	0.93
CaT-B505	Imperfect $(CT)_{17}CC(CT)_2$	DQ090013	F: AGAGAACGACTTTGTATGACAAAGA R: TTGAACCATTAAATACATCATGTGA	55	110–130	6	0.80	0.90	−0.057	0.77	0.87
CaT-B507	Imperfect $(GA)_1GC(GA)_2GC(GA)_{14}$	DQ090014	F: CTAAGCTCACCAAGAGGAAGTTGAT R: GCTTCTGGGTCTCCTGCTCA	55	182–202	10	0.83	0.90	−0.036	0.81	0.93
CaT-B508	Perfect $(GA)_{10}$	DQ090015	F: GGGTCAAGATTTGATAAAGTGGGA R: GCACTCCACTTGTGCGTTTTC	55	148–170	10	0.67	0.65	0.013	0.65	0.84
CaT-B509	Perfect $(GA)_{14}$	DQ090023	F: GTCTGGCATGGTTTTGAGAAGA R: CTTTCCCGCCCAAACCAC	55	109–119	6	0.76	0.85	−0.053	0.72	0.88
CaT-B511	Perfect $(CT)_{18}$	DQ090016	F: CCCCACACTTCCGAATTTAC R: CACACGTTGGAGAATGGTGGT	55	139–158	9	0.82	1.00	−0.100	0.79	0.87
CaT-B512	Perfect $(CT)_{14}$	DQ090017	F: GTGGGCTTTTTAGTTTTTAGCC R: GAAGTTGGAGTAGGGTGATGAG	55	130–151	4	0.31	0.30	0.007	0.29	0.49
CaT-C001	Perfect $(CT)_{15}$	DQ090018	F: AGCTCTTCGTCGTCCGTGAC R: CACAATCCGACACCTACCATC	55	156–175	7	0.77	0.80	−0.014	0.74	0.89
CaT-C502	Imperfect $(CTT)_1T_2(CTT)_{11}C_2T_4CTT$	DQ090019	F: GCATGCAAGGTGGTCGGT R: TTTGGCACCCAACAACTCTAGA	55	151–166	4	0.57	0.50	0.042	0.47	0.74
CaT-C503	Perfect $(GAA)_5$	DQ090020	F: TGCCTCCGCAACGATCAC R: AGCGTCCAGAGAACAATGGG	55	116–119	2	0.05	0.05	−0.001	0.05	0.10
CaT-C504	Imperfect $(CTT)_2T(CTT)_8$	DQ090021	F: GGTCTCCTTCGCTAACATAACCAA R: GTTGCCCTCGAGTTGTAGTA	55	155–176	6	0.76	0.80	−0.023	0.73	0.86
CaT-C505	Imperfect $(TTC)_4T_3GT(TTC)_4C(TTC)_3$	DQ090022	F: TTGCGCTTTGATCAGGACGA R: GACCACCTGCCATTCTGCTA	55	222–225	2	0.05	0.05	−0.001	0.05	0.10

The transportability of microsatellite loci was evaluated in seven *Corylus* species. All tested primer pairs amplified DNA fragments in all species, with two exceptions. CaT-C505 failed to amplify in *C. ferox* and CaT-A114 failed to amplify in *C. californica*. A locus was considered conserved when one or two sharp bands of the expected size were obtained from PCR. Fifteen microsatellite loci produced amplified products of the expected length in all species tested. CaT-C505 was successful in all *Corylus* spp. but yielded a 346 bp fragment in *C. californica*. CaT-C502 showed the presence of multiple bands in *C. chinensis*, *C. colurna* and *C. papyraceae* as did CaT-B512 in *C. americana*, *C. heterophylla* and *C. papyraceae*.

These microsatellite markers will be used for the genetic characterization of hazelnut germplasm and to saturate a linkage map under construction by the OSU hazelnut breeding program.

Acknowledgements

This research was supported by the Foundation Cassa di Risparmio di Cuneo, Italy.

References

Botstein JE, White RL, Skolnick M, Davis R (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics*, **32**, 314–331.

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, **5**, 453–455.

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. *BioTechiques*, **20**, 758–760.

Kasapligil B (1972) A bibliography on *Corylus* (Betulaceae) with annotations. *Annual Report of the Northern Nut Growers Association*, **63**, 107–162.

Kloosterman AD, Budowle B, Daselaar P (1993) PCR-amplification and detection of the human DIS80 VNTR locus. Amplification conditions, population genetics and application in forensic analysis. *International Journal of Legal Medicine*, **105**, 257–264.

Mehlenbacher SA (1991) Hazelnuts (*Corylus*). Genetic resources of temperate fruit and nut crops. *Acta Horticulturae*, **290**, 791–836.

Romisondo P, Me G, Radicati L (1983) Genetic improvement results on hazelnut through breeding and mutagenesis. *Proceedings of the International Hazelnut Congress, Avellino (Italy)* **22–24 September**. Regione Campania — Assessorato Agricoltura Caccia, Pesca e Foreste, Napoli, Italy, 231–242.

Wagner HW, Sefc KM (1999) *IDENTITY 1.0*. Centre for Applied Genetics, University Agricultural Sciences, Vienna.